



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2015

Parental preferences for primary school characteristics

Borghans, Lex ; Golsteyn, Bart H H ; Zölitz, Ulf

Abstract: Free school choice has often been argued to be a tide that lifts school quality through increased competition. This paper analyzes the underlying assumption that school quality is an important choice criterion for parents. Using a large and representative data set of over 15,000 Dutch primary school starters, we estimate models of school demand that incorporate heterogeneity in school preferences. Our results show that traditional measures for school quality matter, but other characteristics, such as school denomination and educational philosophy, are more important predictors of choice. Preferences for these school characteristics are strongly heterogeneous across parents.

DOI: <https://doi.org/10.1515/bejeap-2014-0032>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-175336>

Journal Article

Accepted Version

Originally published at:

Borghans, Lex; Golsteyn, Bart H H; Zölitz, Ulf (2015). Parental preferences for primary school characteristics. *Contributions to Economic Analysis Policy*, 15(1):85-117.

DOI: <https://doi.org/10.1515/bejeap-2014-0032>

Parental Preferences for Primary School Characteristics^{*}

Lex Borghans^a

Bart H.H. Golsteyn^b

Ulf Zölitz^c

Abstract

Free school choice has often been argued to be a tide that lifts school quality through increased competition. This paper analyzes the underlying assumption that school quality is an important choice criterion for parents. Using a large and representative data set of over 15,000 Dutch primary school starters we estimate models of school demand that incorporate heterogeneity in school preferences. Our results show that traditional measures for school quality matter, but other characteristics, such as school denomination and educational philosophy, are more important predictors of choice. Preferences for these school characteristics are strongly heterogeneous across parents.

Keywords: School choice, school quality, school competition

JEL classification: I2, I24, J24

^{*} We thank Andries de Grip, Pierre Koning, Tyas Prevoo, Margriet van der Sluis, Benedikt Vogt, two anonymous referees, and conference participants at the ESPE (2013), IWAAE (2013) and ORD (2013) for valuable comments. We further thank Paul Jungbluth for generously sharing data and for providing valuable background information.

^a Department of Economics and Research Centre for Education and the Labour Market (ROA), Maastricht University, P.O. Box 616, 6200 MD, Maastricht, the Netherlands, lex.borghans@maastrichtuniversity.nl.

^b Department of Economics, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, the Netherlands, and SOFI, Stockholm University, Sweden, b.golsteyn@maastrichtuniversity.nl.

^c Corresponding author. Department of Economics, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, the Netherlands, Tel.: +31-43-3882000, u.zoelitz@maastrichtuniversity.nl.

1. Introduction

Policies that aim to raise school quality, like the no child left behind act in the US, focus on increasing the freedom of school choice. The idea behind these policies is that schools will have to provide higher quality education if they have to compete for students with other schools.

One key underlying assumption for school competition to raise the quality of education is that the quality of a school is an important choice criterion for parents. Parental preferences however are unlikely to be one-dimensional. Parents may also value other characteristics of schools such as the teaching philosophy, religious affiliation or geographical proximity. If parents have weak preferences for school quality relative to other school characteristics, quality competition between schools is likely to remain low, even when parents are given more opportunities to choose. In this case, free choice may stimulate schools to specialize by offering specific features instead of improving the quality of their program. Estimates of parental preferences for primary school quality (measured by test scores) are scarce¹ and the relative importance of other school characteristics has mostly been neglected.

This article provides an elaborate picture of parental primary school preferences by investigating heterogeneous preferences for school quality and a variety of additional school characteristics such as school denomination and teaching philosophy. The measures for school quality that we use are the assessment outcome of the Dutch Inspectorate of Education and school average test scores (CITO scores). Our study complements previous research by investigating to which extent school quality and these additional characteristics form preferences for schools and by studying how these preferences vary across parents.²

¹ Important estimates of parental preferences for school quality are provided by among others Burgess *et al.* (2014), Koning and van der Wiel (2013) Hastings, Kane, and Staiger (2010) and Jacob and Lefgren (2007).

² The effects we identify are not necessarily caused by the quality indicators that we investigate, but could also be driven other unobserved quality attributes that strongly

We use a combination of administrative and survey data on over 15,000 Dutch parents of primary school students to estimate conditional and mixed logit models of school choice. The latter model allows estimates to vary across the population of choosers which is central to the identification of heterogeneity in preferences for school characteristics.

The Dutch context is of particular interest since, in contrast to other countries that have introduced free school choice more recently, it provides a case where school choice has traditionally been unrestricted. The observed patterns in school choice can be assumed to represent a general equilibrium and behavior will not be driven by recent policy changes or other interventions.

Our results firstly show that school quality matters: parents prefer schools with higher average student test scores and schools with a positive quality assessment of the Dutch Inspectorate of Education. We find heterogeneity in the response to the outcome of the Inspectorate's assessment. Parents with higher education levels are more sensitive to the Inspectorate's assessment than lower educated families. Similarly, schools with higher student achievements are on average more popular, but there is substantial variation in this preference among the group of lower educated people. About 32 percent of the lower educated families prefer schools with lower average achievement test scores over higher achieving schools. There are two potential explanations for these findings: (1) parents are unaware of quality differences at the school level or (2) they put a low weight on characteristics related to school quality when choosing a school.³

correlate. We therefore aim to investigate whether they are consistent with observable quality indicators.

³ It is important to keep in mind that the results are subject to reverse causality. The school average test scores and the Inspectorates assessment themselves could also be affected by the type of socio-economic subgroups that select into specific schools. In addition to this, the difference in preferences between higher and lower educated families may also be caused by the correlation of academic characteristics with school attributes that are unobservable. Another potential problem may be geographical (or neighborhood) segregation. Higher educated parents may be more likely to move into neighborhoods with higher quality schools.

Secondly, our results show that the schools' denomination and the practice of an alternative teaching philosophy are important school characteristics that matter for choice. Parental preferences are also heterogeneous with regard to these other school characteristics such as whether a school is public, Catholic, Islamic, is dedicated to Montessori education or follows the educational philosophy of Rudolph Steiner. Even within specified socio-economic subgroups of the population, parents have varying preferences for particular school attributes. While some parents are willing to accept long daily travel distances to a school with a particular specialized profile, other parents avoid this school although it is in their immediate vicinity. Examples of these specialized schools are Islamic schools, which are very unpopular for the largest part of the population but very attractive to about 10 percent of the population. We further show that parents prefer schools that match their own religious beliefs. This effect is particularly strong for Islamic parents.

When looking at preferences for school instruction styles we find that non-mainstream schools with an alternative teaching philosophy are on average less popular, but some higher educated parents strongly prefer these schools. Our estimates show that 26 percent of the higher educated parents prefer schools with alternative teaching concepts over mainstream education and are willing to accept a three to four times longer daily traveling distance to attend such a school.

Our findings with respect to the preferences for school quality are in line with previous research that has shown that parents with a lower socio-economic background put lower weights on observable school quality indicators. Hastings *et al.* (2010) use data from Charlotte, North Carolina to show that black and lower income families have weaker preferences for primary and secondary schools with higher student achievements. The authors conclude that lower quality schools have weak incentives to improve because they attract parents that care less about quality. The authors further argue that high quality schools

perceive more competitive pressure, because the higher SES parents that choose these schools are more sensitive to differences in average student achievements.

Our findings are consistent with the recent work of Burgess *et al.* (2014) who show that parents in England value academic performance and proximity. Our results confirm their finding that socio-economically advantaged parents have stronger preferences for academic performance measured by test scores.

This article is also related to research that has focused on estimating parental preferences for secondary schools. Koning and van der Wiel (2013) investigate how parents in the Netherlands react to publicly available quality information about secondary schools.⁴ They show that newspaper published school rankings significantly predict future student enrolment: negative school quality scores decrease the number of enrolments and positive scores increase the number of students choosing that school. The largest effect was found in response to scores about the highest secondary school track (*VWO*). However, the size of the estimates suggests that even if parents are well informed, the effect of published quality indicators is generally small compared to the impact of other characteristics such as the overall reputation of a school track or the distance from home to school.

We show that quality indicators such as school average achievement test scores and the online published assessment outcome of the Inspectorate of Education matter for primary school choice. Besides school quality, the school's denomination, teaching concept and proximity are important determinants for primary school choice. Preferences are strongly heterogeneous across parents.

⁴ Ruijs and Oosterbeek (2012) also investigate secondary school choice behavior in the city of Amsterdam and conclude that proximity to schools and peers are more important determinants of school choice than quality.

The remainder of this article is structured as follows. Section 2 provides background information on the Dutch education system. Section 3 discusses the data. Section 4 describes our empirical strategy. Section 5 shows the estimation results. Section 6 concludes.

2. The Dutch education system

The Netherlands have a longstanding tradition in free school choice. Since 1917, parents can freely choose between state funded private and public schools. In contrast to other countries where free school choice policies are introduced more recently, the Netherlands provide a case to study school choice behavior that can be assumed to be closer to a “steady-state”. The results presented in this article can therefore be interpreted as evidence about how parental school choice preferences are shaped in a stable system without major policy discontinuities. This can be of particular interest for policy makers who are interested in the persistence of observed short run effects of free choice.

Dutch parents can in theory choose among all schools in the Netherlands. There are no school catchment areas or choice restricting school districts. School tuition fees are small or non-existent and therefore do not restrict school choice. The most obvious costs that parents face when choosing a more distant school are the opportunity costs of time. Both parents and their children will spend more time on traveling if they choose a more distant school since the school and the classmates of their children will be located farther away. Due to the high population density in the Netherlands, a number of schools is usually within walking or cycling distance.⁵

Dutch parents can choose among schools with a variety of different religious affiliations and diverse educational concepts that fit their personal values best. Besides public schools (“openbare scholen”) there are schools with denominations like Catholic, Hindu,

⁵ The population density is about 487 people per square kilometer (if only the land area is being counted).

Islamic, and Protestant schools. Schools which focus on a specific pedagogical concept are the “Brede school”, “Dalton school”, “Freinetonderwijs”, “Iederwijs”, “Jenaplan”, “Montessori” and “Vrije scholen”. The three most prevalent schools types that are usually within reach are Catholic, Protestant and public schools.

All school types are publicly funded, proportional to their student numbers. Schools that serve students from a more disadvantaged parental background receive some additional funding.

Oversubscriptions are sometimes a problem in larger cities in the western part of the country and are usually solved with admission lotteries. In some urban areas of the Netherlands, especially Amsterdam and Utrecht, student numbers are rapidly growing and white parents with a higher socio-economics background seem to prefer schools with a high fraction of parents with a similar background. Therefore a number of “white” schools in these areas are heavily oversubscribed. In the most southern province Limburg, where the data for this article were collected, there is little segregation. The area faced a decline in the number of children in recent years. Therefore, oversubscriptions and capacity constraints do not play a role for choice considerations.

2.1 Primary education

Primary education in the Netherlands starts at the age of four and is compulsory from the age of 5. Parents in the region South Limburg apply directly at the schools where they would like to enroll their child. According to Dutch law, all schools (with a few exceptions) are obliged to accept all children that wish to enroll unless applications exceed capacity constraints. In the region that we analyze most children are accepted at the first school they apply.⁶ A large share

⁶ In the latest survey which was sent out to the parents, the following question was included in a section about the school which the child is attending: “Was this school your first choice?”.

of the schools in the Netherlands has a religious affiliation. For these schools, the same public finding rules apply as for public schools. In principle these schools could refuse children that do not practice their religious principles, but schools are required to behave consistently in this respect. That means that once one child is accepted that does not meet these norms, the norm cannot be applied anymore. There are schools in other regions in the Netherlands that have access restrictions for these reasons, but in the region we analyze there are no schools that apply these specific norms.

Primary education in the Netherlands lasts for eight years, where the first two grades consist mostly of preparatory activities and can be compared to kindergarten in the US. At the end of grade eight, a nationwide standardized central exam is taken – the so-called CITO test. This test is externally developed and objectively marked by the CITO institute, a non-profit organization, which is independent from the schools. So the test is not assessed by teachers of a particular school themselves. The conduction of the test takes place at the same time for all schools and lasts for several school days. The CITO test covers three main subjects: language, math and science. The outcome of the CITO test together with the teacher's recommendation determines the secondary school track. There are three different types of secondary education tracks that vary in their degree of difficulty and educational focus. Only the highest track (VWO) permits access to studies at the University level. For the purpose of this paper use the school average performance of students in grade 8 to predict the choice outcomes of children that just entered in grade 1. Therefore CITO scores are likely to contain latent lagged measures of school quality of the past eight years.

Figure 1 shows the distribution of the end of primary school student achievement tests and the recommended thresholds for secondary school tracking. The figure shows that the

93.8 percent of the parents state that this is the case. This provides evidence that oversubscriptions are not a serious problem in this part of the Netherlands.

variation in the average achievement test scores across schools is substantial.⁷ If a school is to the right of the solid vertical line, the average student at this school goes to the highest secondary school track. The average achievement test score of a school is generally not publicly known, but some schools publish their score on their websites for marketing reasons.

2.2 Inspectorate of Education

In order to maintain certain standards at the school level, the educational quality is monitored by the Dutch Inspectorate of Education which assesses schools through regular visitations and by evaluating student achievements. The Inspectorate's assessment consists of class observations, and meetings with pupils, parents, teachers and the school management team. In addition, the Inspectorate investigates whether the school is able to perform a reliable self-evaluation and whether the school complies with statutory rules and regulations. One additional factor in the process of evaluation is the schools' average achievement test score (CITO score).⁸ There are three final outcomes of these quality assessments: "trusted", "weak" or "unsatisfactory". About five percent of the primary schools receive the label "weak" and less than 1 percent are "unsatisfactory". The quality assessment of the Inspectorate is published online and parents can look up the rating of a specific school or compare the schools in the neighborhood. If schools continue to provide unsatisfactory quality or do not comply with the Inspectorate, they can be sanctioned by the Minister of Education. The most extreme measure is cutting off a school's funding.

3. Data

⁷ One interesting question from an economic perspective is why these differences in school quality maintain in a system of free choice. Market forces should lead to an equilibrium where all schools converge to a certain quality level. A potential answer could lie in the heterogeneity of preferences.

⁸ About 90 percent of all schools use the "CITO" test for student assessment.

The school choice data analyzed in this article were collected by the Maastricht University research project *Moelejaan* in the southern part of the Dutch province Limburg. South Limburg is an urbanized area of 661 square kilometers and a population of about 620,000. It is surrounded by Belgium and Germany and only the north is connected to the rest of the Netherlands.

The data we use are a combination of administrative student data from schools and survey data from a questionnaire among the parents at these schools. 200 primary schools in this area participated in the study and provided student level data. This is about 97 percent of the primary schools in the region South Limburg. The data consist of 16,852 children from three different cohorts that started school in 2007, 2008 and 2009. For the analysis, 1,065 children are dropped from the sample due to missing or non-existent home addresses. 13 schools are excluded because of missing end of primary school achievement test scores (CITO scores).⁹ The final sample consists of 15,000 students from 183 schools. For about 64 percent of these children we have information about the education level of the parents from survey data, which we use for some parts of our analysis.

Table 1 shows descriptive statistics for five different school types. The table gives a first impression of how parents at various schools differ in terms of observables. Column (5) shows that schools with an alternative teaching concept attract particular students. Students at these schools travel a substantially longer distance from home to school and have on average higher educated parents. The performance of these schools is relatively weak in terms of student achievement and the schools are also more likely to be assessed as “weak” by the Inspectorate.

The table also shows the average distance rank of the students that attend a particular school type. The average distance rank of all schools is 2.66, which means that students on

⁹ These schools did not participate in the CITO test. They conducted a non-comparable end of primary education achievement tests.

average have 1.66 school options that are closer to their home location, but which they do not choose to attend. For schools with an alternative teaching concept, the average student-school distance rank is 5.78. The table shows that more specialized schools attract students which live farther away (i.e. that have more other schools closer to their homes).

Table 2 shows that there are many schools in the direct vicinity of the students' homes. Figure 2 shows the distribution of the Euclidian distance to the chosen school. The median distance to the chosen school is .525 kilometers. 90 percent of the parents choose a school that is within a 1.785 kilometer radius from home, and few parents choose schools located far away from their home.

4. Empirical strategy

In order to investigate how parents choose between different primary schools we apply two different discrete choice models: the conditional logit and the mixed logit model. In the following we will briefly discuss the intuition and limitations of both estimation techniques.

We use both the conditional and mixed logit models to allow a direct comparison between the different methods. Our benchmark model in the analysis is the mixed logit model. This model is less restrictive than the conditional logit model regarding the independence of irrelevant alternatives (IIA) assumption and crucial to identify heterogeneity in preferences without using one specific variable that represents the dimension of heterogeneity. The downside of the mixed logit model is that its lack of a closed form solution creates computational limitations. The mixed logit further requires making an appropriate assumption about the mixing distribution of the random variables. In the following we discuss the functional form of both models in more detail.

4.1 The conditional logit model

The conditional choice model was developed by McFadden (1973) and first applied to the context of school choice by Glazerman (1998). Intuitively, the conditional logit model compares the characteristics of the chosen school with the characteristics of all schools which are not chosen, using student fixed effects.

In this model, the parents of each child i face the decision to choose between j different schools. Every school is associated with a certain amount of individual specific utility U_{ij} . The model assumes that utility U_{ij} can be described as a linear function of a child specific component, a school specific component X_j , a school-child specific component Z_{ij} and a non-systematic random component ε_{ij}

$$U_{ij} = W_i + X_j + Z_{ij} + \varepsilon_{ij} . \quad (1)$$

Examples of W_i are individual specific characteristics like gender, age or ability. X_j are school characteristics that are common to the population of school choosers. Examples of X_j are the average test results of a school, the assessment of the Inspectorate or the denomination of a school. The characteristics Z_{ij} are school-student specific and represent an interaction between child and school attributes. An example of Z_{ij} that we include in our model is the distance between the parents' home and the school.

The utility that the parents of child i obtain from choosing school alternative j can be written in a linear form as:

$$U_{ij}(X_j, Z_{ij}) = \beta_1 W_i + \beta_2 X_j + \beta_3 Z_{ij} + \varepsilon_{ij} \quad (2)$$

Since we estimate preferences for schools within individuals we cannot identify β_1 . In our estimation, all individual characteristics that do not vary over school alternatives are

collinear with the individual fixed effect and therefore excluded from the model. The parameters β_2 and β_3 are the focus of our interest, since they determine how differences in the attributes translate into utility differences. We assume that parents choose the school that provides them with the highest level of utility. The individual specific error terms are assumed to be random, independently-distributed variables with an extreme value distribution (the Gumbel distribution).

We do not directly observe the utility level U_{ij} , but use the observed outcomes of the choice process which are revealed to us in the data. If the parents of a child choose school J , it is revealed to us that $U_{iJ} > U_{ij-J}$, i.e. that the utility level of the chosen school must have been larger than the utility of any other school in the choice set. The observed outcome variable Y_{ij} is binary and takes on the value 1 if the school is chosen and zero otherwise. Using the observed outcomes of school choice we can estimate how differences in X_j and Z_{ij} , which vary over different school options, relate to the probability that one given school has the highest amount of utility of all schools in the choice set.

W_i , the individual specific characteristics do not vary over different choice alternatives. When estimating the model, all individual characteristics are invariant and fully collinear with the included individual fixed effects. Therefore W_i do not affect the choice probability. The resulting conditional logit model analyzes within person differences in preferences for schools. At first glance this appears unattractive since one of the main research questions of this article is how preferences for certain school characteristics vary across different socio-economic subgroups. In the conditional logit framework, estimating heterogeneity with respect to individual characteristics can be elicited by (1) estimating the model for subsamples of the population separately, or (2) including interaction terms between individual characteristics and school specific attributes.

One practical problem with the conditional logit estimation is the assumption of independence of irrelevant alternative (IIA).¹⁰ The IIA property assumes that the probability of choosing a particular alternative does not depend on the characteristics of other alternatives. Mixed logit models do not rely on this assumption.

4.2 The mixed logit model

The mixed logit model allows estimating the mean preference for a school attribute and the standard deviation of the mean preference coefficient. This appears particularly attractive for our research question since it allows a more direct estimation of heterogeneous preferences for certain school characteristics. The main advantage relative to the conditional logit model is that we do not need to specify the dimension in which a preference is heterogeneous. Using the mean and standard deviation, we can directly infer which share of the population has particular preferences for a specific school characteristic.

The mixed logit is also a convenient solution for avoiding the IIA property of the conditional logit. The mixed logit or random parameter logit that we apply was developed by Train (2009) and allows for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors. The mixed logit model works very similarly to the conditional logit model, but does not only estimate one point estimator that represents the mean preference of the population. Instead, the model assumes that coefficients β_i vary over the population and are drawn from a normal or lognormal distribution. The mixed logit model can be represented in the following way in a random utility framework:

¹⁰ To illustrate the importance of this, suppose parents have the choice between Catholic school A, Catholic school B and Protestant school C. The IIA assumption is that parents who have initially chosen school A would be equally likely to switch to school B and C if their chosen school A is closed and everything else including all school characteristics is held constant. One can expect that the IIA assumption does not hold in this school choice context since we would think that parents who have initially chosen for a Catholic school will be more likely to choose a school that is also Catholic.

$$U_{ij} = \beta_i' X_{ij} + \eta_{ij} + \varepsilon_{ij} . \quad (3)$$

The utility level the individual i obtains from choosing school alternative j is described as a function of X_{ij} which is a vector of school and/or child specific characteristics, the random parameter η_{ij} and the “regular” error term ε_{ij} . ε_{ij} has a mean zero and is independent and identically distributed over alternatives.¹¹ It is unrelated to η_{ij} , the parameters in the vector β_i . The density function of η is given as $f(\eta | \beta)$ and depends on the parameters of the model. Following Train (2009), the estimated mixed logit probability is a weighted average of the logit equation, but evaluated at different values of β , with the weights given by the density $f(\eta | \beta)$. For a given fixed β the probability of choosing alternative k , which inherits the highest utility out of J different options is:

$$L_{ik}(\beta_i) = \frac{\exp(x\beta_i' X_{ik})}{\exp(\sum_j \beta_i' X_{ij})} . \quad (4)$$

Since β is not fixed we have to integrate $L_{ik}(\beta_i)$ over all possible variables of β_i . For this, we need the mixing distribution $f(\beta)$. The mixed logit probability is then given as:

$$P_{ik} = \int \left(\frac{\exp(x\beta_i' X_{ik})}{\exp(\sum_j \beta_i' X_{ij})} \right) f(\beta) d\beta \quad (5)$$

¹¹ Further, it is assumed to be not (auto) correlated and to have the same mean and same variance (homoskedastic).

Equation (5) does not have a closed form solution and the probability P_{ik} can be estimated using simulated maximum likelihood.¹²

4.3 Accounting for school composition

For the identification of parental preferences for school characteristics it is important to keep in mind that other school characteristics that correlate with our measures of school quality might be confounding factors. One obvious candidate for such a characteristic could be the fraction of peers with a similar socio-economic background at a given school. However, including the percentage of peers at a given school as an additional school characteristic introduces potentially severe endogeneity problems that prohibit a direct interpretation of peer preference estimates.¹³ One additional empirical challenge for the identification is that achievement tests are likely to reflect school quality but might also be based on *a priori* differences between children. For instance, children might sort into schools based on their socio-economic background.¹⁴

5. Results

¹² For the simulated maximum likelihood estimation one needs to generate random draws from $f(\beta)$. For every individual we use 50 draws of β . Our obtained results were not sensitive to an increase in the number of draws that we use.

¹³ In a robustness analysis, we include the lagged fraction of peers with a similar education level (that have chosen the same school) as an additional school attribute to test to which degree peer preferences matter. These analyses show that the results remain robust to the inclusion of lagged peer characteristics. Due to the inherent endogeneity of this approach that complicates the interpretation of the obtained estimates we do not report these models in this article.

¹⁴ We have also tested the robustness of our estimates using important parental background characteristics available in the data (education level and household income). We run regressions of the raw CITO score on parental characteristics and used the predicted scores of these regressions as a measure of socio-economic background adjusted CITO scores. The adjusted measure of school quality takes into account school level differences in the education level of the parents and the household income of parents. The point estimates of the adjusted CITO are marginally smaller, but this difference is not statistically significant.

We analyze the preference for distance in section 5.1 and the preference for school quality in section 5.2. Section 5.3 analyzes the heterogeneity of preferences with respect to these indicators and section 5.4 investigates to what extent school popularity varies between socio-economic groups. Section 5.5 and section 5.6 respectively analyze the importance of school denomination and alternative educational concepts.

5.1 Preferences for distance

Table 3 shows the estimates we obtain from the conditional logit estimation. The displayed coefficients represent the raw coefficients. Coefficients significantly larger than zero represent a characteristic that is related to an increase in the probability of choice and coefficients significantly lower than zero mean that this school characteristic is unpopular or unattractive among the population of choosers. The table shows that increased distance to a school option is negatively related to the probability of choice. Parents appear to have a strong preference for choosing a school close to their home location.

Figure 3 shows how the probability of choosing a school relates to the distance of a school option and investigates whether this effect is linear. The graph is based on a conditional logit estimation that includes one dummy variable for every 100 meter unit of distance between home and school. The omitted category is having a school option within 100 meters from home. The vertical axis displays the size of the coefficients and their 95 percent confidence intervals. As the distance between the home address and the school increases, the likelihood of choice decreases in a slightly convex trend. The difference between 0 and 1 kilometer reduces the probability of choice more than the difference between 3 and 4 kilometers. The marginal costs of traveling distance decrease slightly with increased distance.

5.2 Preferences for school quality

The regressions in table 3 also include two measures for school quality: the outcome of the assessment of the Inspectorate, which is published online and the three year school average achievement test score (CITO score), which is not available for all schools. A negative assessment by the Inspectorate of Education is negatively related to the choice probability. Higher average student achievement test scores relate positively to the choice probabilities. These findings support the hypothesis that parents value these two measures of school quality when choosing among the different schools. In relation to the strong effect of distance, the preference for school quality seems to be rather weak. According to the estimated model, a school with a “trusted” rather than “weak” assessment outcome increases the willingness to travel farther by 190 meters.¹⁵ The inclusion of quality indicators does not attenuate the point estimates for distance. This shows that parents’ preferences for distance appear to be uncorrelated with academic quality (as measured by the CITO score or Inspection outcome).

The estimates that we obtain from estimating the mixed logit models in table 4 show similar results. In table 4 we report the mean of the logit estimator and the standard deviation of the respective random parameter. The model shows that there is significant variation in the distance parameter, which suggests that some subgroups of the population are more sensitive to an increase in traveling distance than others.

In both types of models, the size of the “weak assessment” coefficient is smaller when the standardized achievement test score is included. This collinearity between the two quality indicators is due to the construction of the Inspectorate’s assessment, which is partly built on the achievement test score (CITO).

We also investigate whether the results may to some extent be driven by selective response to the survey. Table A1 shows which variables are correlated to survey response.

¹⁵ The implied willingness to travel can be obtained by dividing the ‘Weak’ assessment coefficient by the distance coefficient in table 3. Differently framed example: Parents prefer school A with a CITO average of 530 over school B with a CITO average of 540 that is 100m farther away than school A.

Responders are on average more likely to live closer to the chosen school, less likely to attend a school with a “weak” inspectorate assessment and more likely to come from a school with a higher CITO score. Table A2 shows how our point estimates differ between the full and the questionnaire sample. The point estimates are somewhat larger in the questionnaire sample, but they do not differ in a qualitative sense. In order to look at how preferences differ between socio-economic subgroups we are restricted to use the questionnaire since we do not have information about the parental education level for the rest of the sample.

5.3 Heterogeneity across socio-economic subgroups

We now turn to the question whether preferences for school quality are heterogeneous among parents with a different educational background. In order to investigate heterogeneity, we split the sample based on the highest obtained educational degree of the father of the child.¹⁶ Table 5 shows estimates obtained from the conditional logit estimation for the three different subgroups. The estimates indicate that preferences for school quality vary substantially across groups. Higher educated families have stronger preferences for school average achievement test scores (CITO) when compared to middle and lower educated families. These last two groups do not significantly differ in their “taste” for high student achievements.

The interpretation of these coefficients requires some caution. What appears to be a difference in preferences could also be a result of asymmetric information about school quality. The achievement test (CITO) scores may not be known to the public as only some schools publish them on their websites. Lower educated families may in general have less information about school quality or face larger information costs. Besides that, schools with

¹⁶ The group “higher educated” refers to families where the father holds a degree from higher tertiary education: professional bachelor education (HBO) or university. The middle educated group refers to a degree from high secondary education (VWO) as the highest obtained degree of the father and the “lower educated group” represents families where the father holds a lower secondary or tertiary education level (HAVO/(V)MBO) or no educational degree.

lower average student achievements may be more likely to be located in neighborhoods with more people from a lower socio-economic background. Our findings therefore do not necessarily have a causal interpretation.

Table 5 further shows that a “weak” outcome of the Inspectorate’s assessment is negatively related to the probability to choose the school. This effect varies across the three subgroups. Preferences for avoiding “weak” schools increase with the education level of the father. Holding everything else constant, higher educated parents would be willing to accept a 375 meters longer way to school if they thereby could avoid sending their child to a “weak” labeled school. For lower educated families the respective distance is only 127 meters.¹⁷

The preferences for school distance appear to be homogeneous across subgroups. In the estimates that represent the preferences for school proximity (distance to school), we observe only minor differences.¹⁸

The results we obtain from the mixed logit models in table 6 give similar results with respect to the heterogeneity of preferences for school quality across families with different education levels. The benefit of the mixed logit model is that it allows us to investigate whether preferences are homogeneous *within* the specified educational subgroups. Table 6 shows that there is substantial variation in the preference for “weak” schools among the group of lower educated. About 14 percent of the lower educated appear to prefer weak schools over non-weak schools.¹⁹

This finding suggests that a substantial fraction of this subgroup either does not know about the online published Inspectorate’s assessment or that they do not care about what the Inspectorate considers poor educational quality. Yet another explanation could be that this

¹⁷ The implied willingness to travel can be obtained by dividing the ‘Weak’ assessment coefficient by the distance coefficient in table 5.

¹⁸ These differences are small, but significantly different across the three socio-economic subgroups.

¹⁹ This percentage for a given variable X can be obtained by applying the formula: $1 - \Omega(\widehat{\beta}_x / \widehat{SD}_x)$; where Ω describes the cumulative standard normal distribution.

subgroup puts more weight on other school characteristics that negatively correlate with the Inspectorate's assessment. One obvious candidate for such a characteristic would be the fraction of peers with a similar socio-economic background at a given school. Including the percentage of peers at a given school as an additional school characteristic introduces potentially severe endogeneity problems and the interpretation of peer preference estimates is not straightforward.

Figure 4 displays the relationship between school popularity and school average student achievements in a different way. Controlling for distance to the school, we estimate a conditional logit model with a specific dummy variable for each school. The size of the estimated dummy coefficients can be interpreted as its relative popularity relative to other schools given their geographical characteristics. Figure 4 plots the coefficient size of these school dummies against the achievement test (CITO) score of the respective school. The vertical axis represents the size of the school dummies from a conditional logit model. The horizontal axis displays the school average CITO score of the respective school. The slope of the fitted values line is positive and significantly different from zero, which means that the school's "attractiveness" relates positively to the average student achievements of the school. However, it is apparent from the graph that student achievements alone are rather weak or noisy predictors of school popularity.

5.4 Consistency of school popularity across socioeconomic subgroups

In the following we investigate whether different subgroups of the population consider a given school similarly attractive. We estimate a conditional logit model that includes distance to school and one dummy for each school in our data set. We estimate this model for three educational subgroups separately. These groups are constructed based on the highest obtained

degree of the father.²⁰ The estimation output of this model produces 181 school fixed effects for each of the three educational subgroups. Each of these fixed effects yields a point estimate that represents the relative attractiveness in terms of revealed parental preferences.

The attractiveness of a given school in one socio-economic subgroup can now be compared to the revealed preference for this school in a different subgroup by looking at the correlation of the schools' fixed effects between two different groups. Figure 5 shows the consistency of school popularity across educational subgroups. Figure 5a displays the size of the school "attractiveness" dummy for the higher educated on the vertical axis and the size of the same school dummy for middle educated on the horizontal axis. With fully homogeneous preferences for schools the slope of fitted values should be identical to a 45 degree line. The slope of the fitted value line is .93 and not statistically different from one. This correlation suggests that high and middle educated parents find the same schools attractive and have very similar revealed preferences. In contrast to that, the correlation in preferences of higher and lower educated families is much less consistent (Figure 5b). The correlation of revealed school attractiveness dummies is only .41 and statistically different from one. In figure 5c, where we compare the revealed preferences of middle and lower educated parents, the slope of the fitted value line is only .36, and statistically significantly different from one. This comparison between middle and lower educated parents shows that preferences for schools are remarkably different between these two groups.

5.5 Preferences for school's denomination

We now investigate whether a school's religious affiliation and educational concept matter for parental choice behavior. The mixed logit models in table 7 shows our most complete model of school choice. As in the previous regressions we include distance and two indicators for

²⁰ We split the sample into higher, middle and lower educated families. Using the education level of the mother provides qualitatively very similar results.

school quality. In addition to that, we also include three dummies for Public, Protestant and Islamic schools. The omitted reference category is Catholic schools.

Column (1) in table 7 shows that public schools appear on average equally attractive as Catholic schools, but there is substantial variation in the preference parameter. About half of the population prefers one of the two types. Looking at the group of middle educated parents we see that about 40 percent prefer Catholic over public schools. Protestant schools have on average a lower probability of being chosen than Catholic schools for higher and middle educated parents. About 15 percent of the higher educated parents prefer Protestant schools over Catholic schools. About 20 percent of the middle educated prefer Protestant schools over Catholic schools.

The one Islamic school in our dataset is on average very unpopular for most parents, but this preference varies substantially over the population as revealed by the estimate of the standard deviation. About 10 percent of the parents have orthogonal preferences and find the Islamic school attractive.

Since the survey data we are using also contain a question on the religious affiliation of the parents, we can test whether preferences for the religious denomination of the school depend on the parents' own denomination. In table 8 we show how preferences for school denomination differ by the religion of parents. Column (1) shows that parents who indicated that they are not religious prefer public schools over Catholic and other religious schools. As expected, column (2) shows that Catholic parents prefer Catholic schools for their children. Column (3) shows that Islamic parents have particularly strong preferences for sending their children to an Islamic school. For this group, preferences are not only strong, but also quite heterogeneous. The majority of Islamic parents seem to strongly prefer Islamic schools while others want to avoid these schools. Taken together the evidence presented in table 8 shows that parents seem to prefer school that match their own religious beliefs.

5.6 Preferences for alternative educational concepts

The regressions in table 7 also include a dummy that specifies whether the school is devoted to an alternative teaching concept. Schools that belong to the “alternative teaching concept” category are Montessori, Jenaplan and Vrije schools. These schools usually grant their students a larger degree of individual freedom during the learning process and put more focus on fine arts, cultural, and artistic education. The reference group in the regressions is schools with a mainstream educational concept. The mean preference shows that schools with an alternative teaching philosophy are generally less popular. However, there is substantial variation in the preference for these schools. About 26 percent of the higher educated parents prefer these schools over mainstream schools. Middle and lower educated families seem to largely avoid these schools.

An important point is that the inclusion of variables for different denomination types and an alternative teaching concept does not systematically alter the preference estimates for achievement test scores and the Inspectorate’s assessment. This supports that parental preferences for these special school characteristics are not captured by the available traditional measures of school quality.

Schools with higher average student attainments and better quality assessments have a higher probability of being chosen, but parental preferences for certain denomination types and an alternative teaching concept are heterogeneous and particularly strong in some subgroups of the population.²¹

6. Conclusion

²¹ For CITO scores there is substantial variation in the preference among the group of lower educated people. About 32 percent of the lower educated families prefer lower CITO score schools over higher CITO score schools.

This article investigates which school characteristics matter for parental primary school choice in the well-established system of free school choice in the Netherlands. We find that school quality (measured by school average test scores) predicts school choice. Publicly available information on school assessments seems to influence choice behavior, particularly of higher educated families. However this “taste for quality” is dominated by the strong preference for schools that are close to the homes of the parents. Our findings confirm evidence from the US, which shows that lower educated parents appear to have weaker preferences for school quality.

Moreover, our findings show that parents have strong preferences for a schools’ religious affiliation and the educational philosophy that is being applied at a school. Specific subgroups of the population have orthogonal preferences for a given school attribute. Parental choice preferences regarding these specialized schools with particular educational concepts have previously not been investigated.

Our results have implications for the discussion about the long run effects of free school choice that currently takes place in the US, UK and many other countries. Some schools in the Netherlands seem to have specialized on particular socio-economic subgroups with specific norms and values. From a school perspective, specialization might be a rational response when being faced with more transparency about school quality and increased competition for students. Policies aiming to increase school competition may be the tide that lifts all boats, but they might also foster school specialization in the long run. The case of the Netherlands shows that specialized schools are able to offer something valuable to parents, that is different from traditionally measured academic quality.

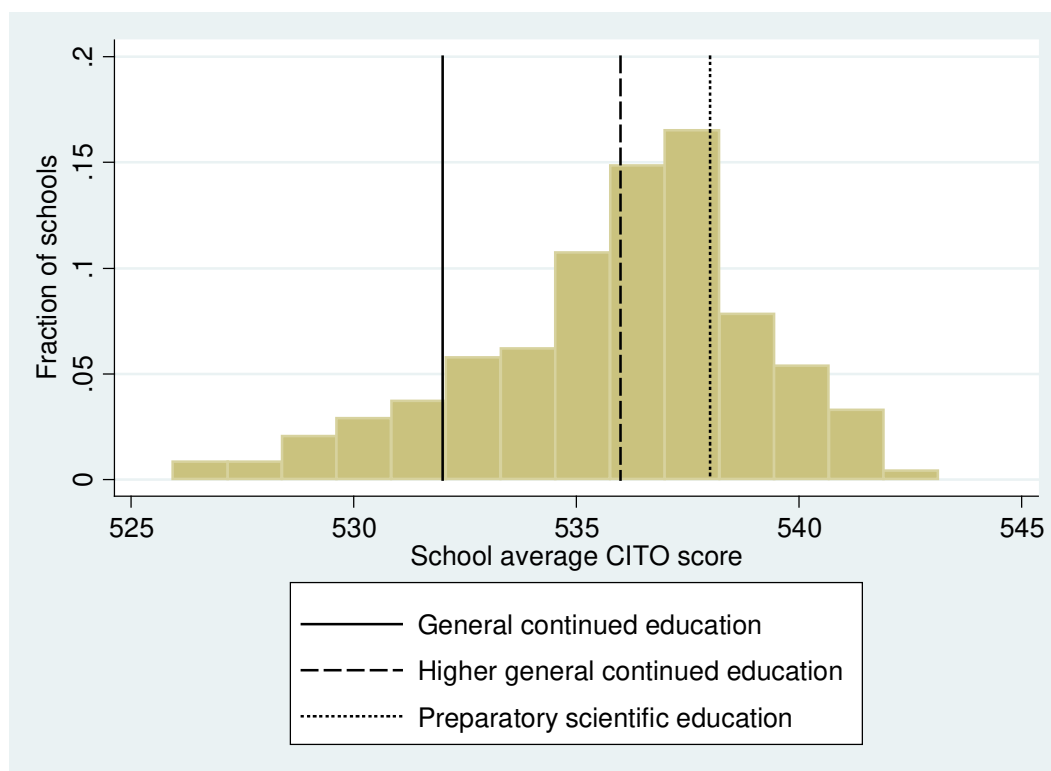
References

- Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2014). "What Parents Want: School Preferences and School Choice." forthcoming in: *The Economic Journal*.
- Glazerman, S. (1998). "School Quality and Social Stratification: The Determinants and Consequences of Parental School Choice." *Unpublished manuscript*.
- Hastings, J., Kane, T. and Staiger, D. (2010). "Heterogeneous Preferences and the Efficacy of Public School Choice." NBER Working Paper Series 2145.
- Hastings, J., and Weinstein, J. (2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics*, 123, 1373–1414.
- Hastings, J. S., Van Weelden, R. and Weinstein, J. (2007). "Preferences, Information and Parental Choice Behavior in Public School Choice." NBER Working Paper Series, 12995.
- Jacob, B., and Lefgren, L. (2007). "What do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers.", *Quarterly Journal of Economics*, 122 (4): 1603-1637.
- Koning, P. and Van der Wiel, K. (2013). "Ranking the Schools: How Quality Information Affects School Choice in the Netherlands." *Journal of the European Economic Association* 11(2) 466-493.
- Koning, P. and Van der Wiel, K. (2012). "School Responsiveness to Quality Rankings: An Empirical Analysis of Secondary Education in the Netherlands." *De Economist* 160(4), 339-355.
- Ruijs, N., Oosterbeek, H. (2012). "School Choice in Amsterdam. Which Schools do Parents Prefer When School Choice is Free?" Unpublished Manuscript, University of Amsterdam.

Train, K. (2009). “*Discrete Choice Methods with Simulation*” (Second Edition). New York:
Cambridge University Press.

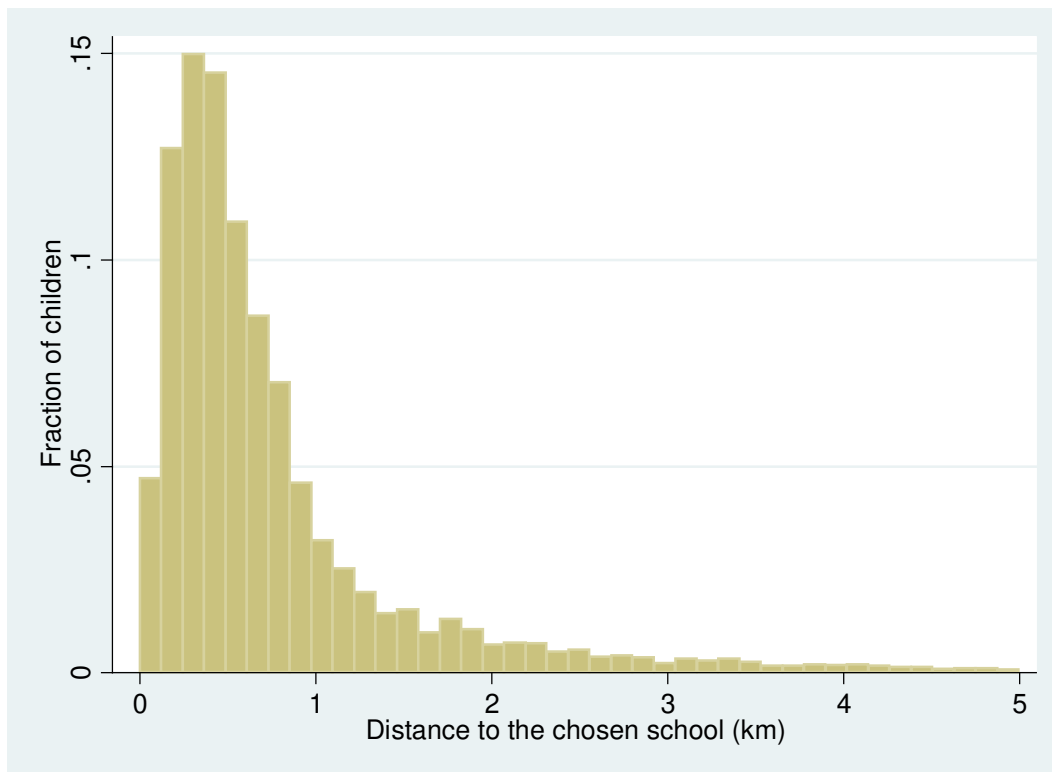
FIGURES

Figure 1: Distribution of school aggregated achievement test score averages



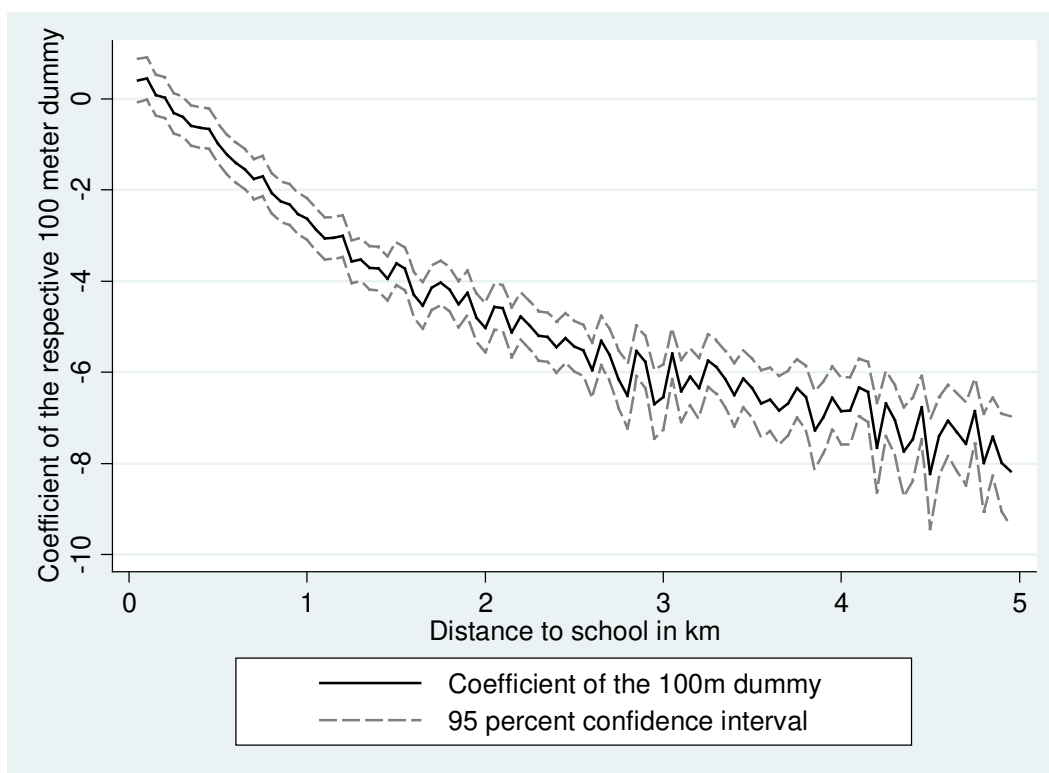
Notes: These are school level averages of national-wide end of primary education achievement tests (CITO scores). The student specific achievement test score ranges between 500 and 550. The vertical lines mark the recommended thresholds for secondary school admission. $N(\text{schools}) = 183$.

Figure 2: Distance to the chosen school in kilometers



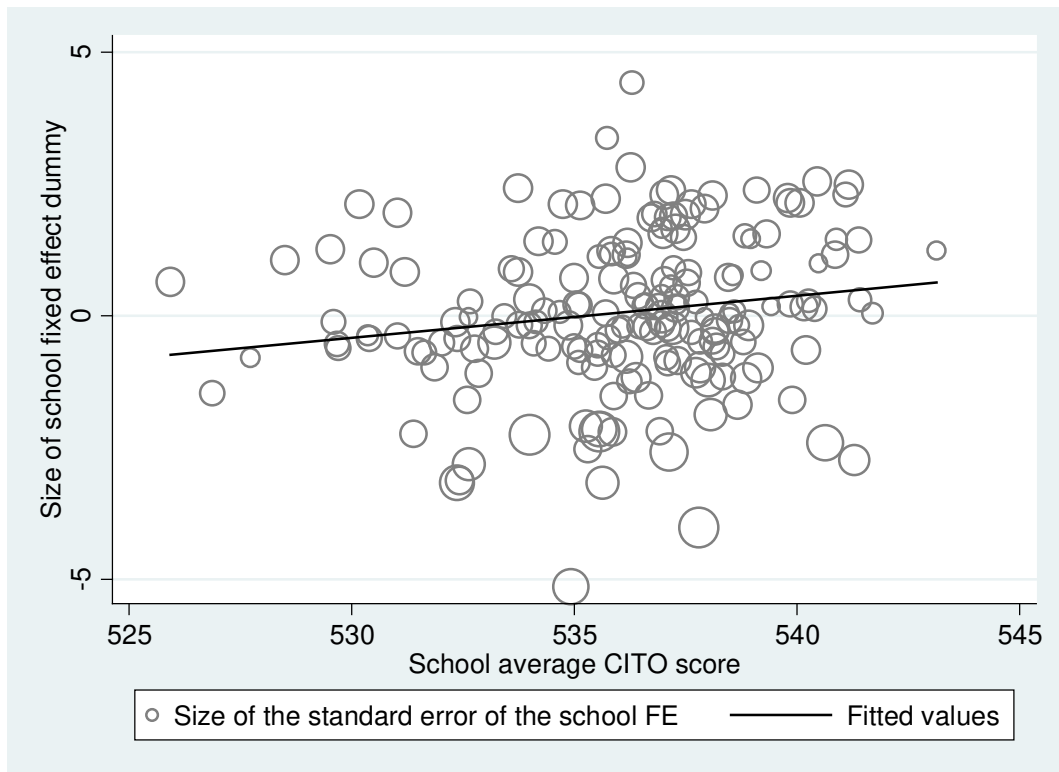
Notes: The figure shows the distribution of the variable “Distance to school”. 131 students attend schools farther than 5 kilometers away. N = 15,435.

Figure 3: Probability of choice and distance



Notes: Development of distance dummies (in 100m steps) from conditional logit estimation and their 95% confidence intervals. Estimations based on 15,304 students from 192 schools. The reference category are school options within 0 - 100 meters distance from home.

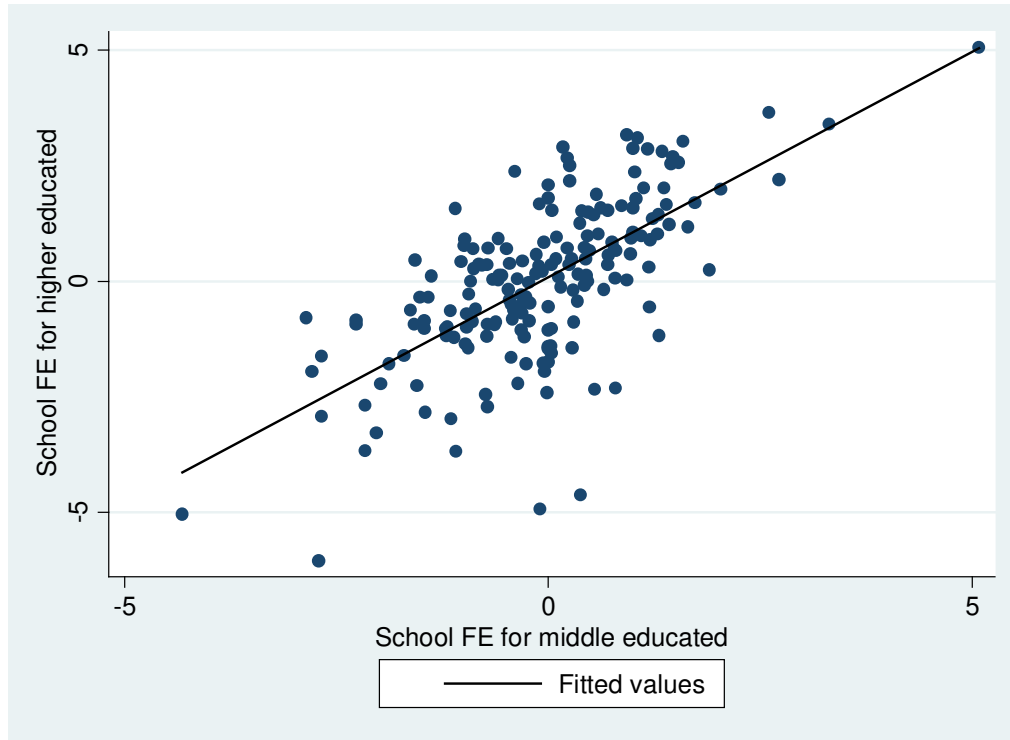
Figure 4: Choice probability and student attainments (CITO)



Notes: The radius of the circle represents the size of school dummy standard error. The slope of the line is significantly positive ($P > |t| = 0.023$). The figure is based on 183 schools. Nine schools were dropped from the sample due to a missing achievement test (CITO) score.

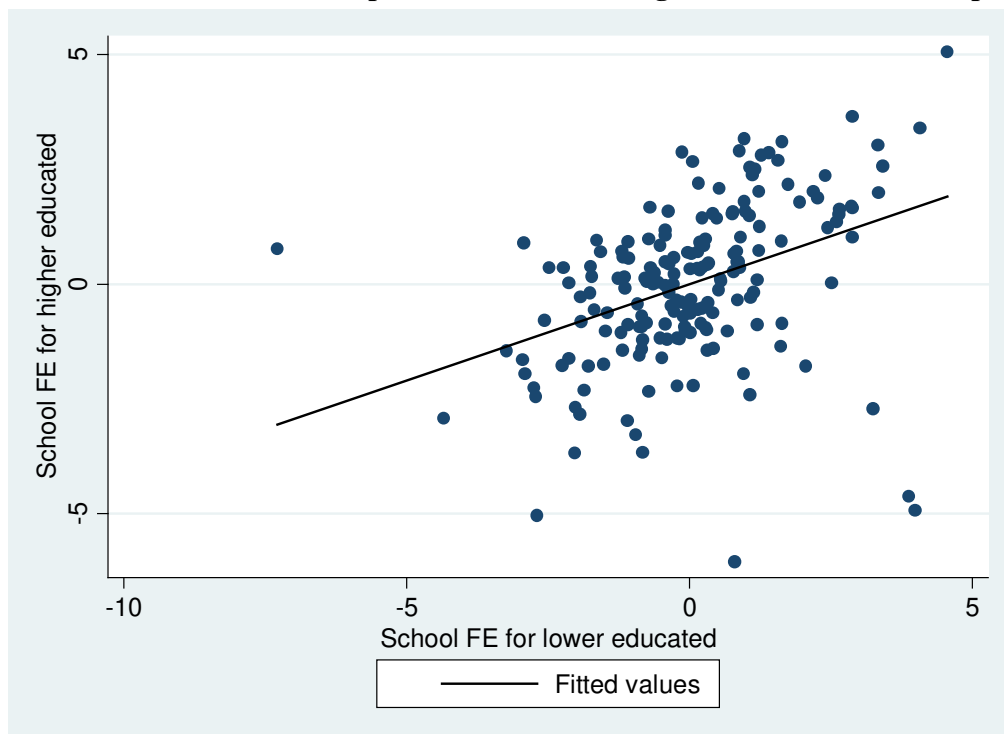
Figure 5: Do different SES subgroups find the same schools attractive?

Figure 5a: Correlation of school preferences between higher and middle educated parents



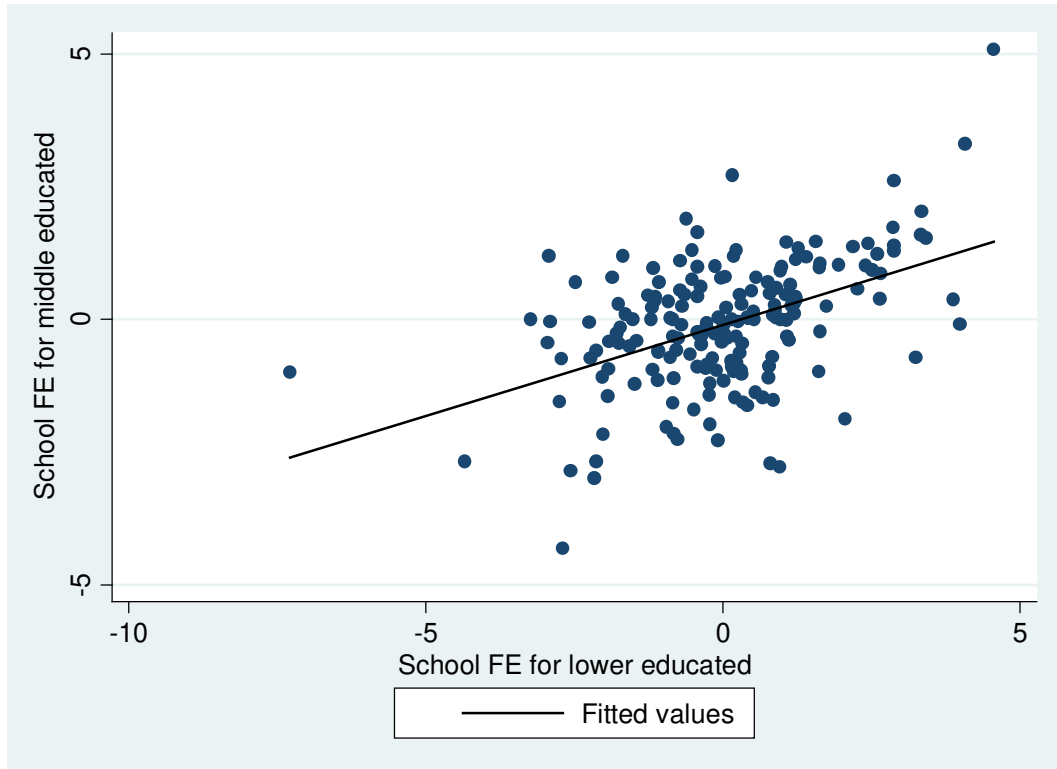
Note: The figure shows how school fixed effects obtained from conditional logit estimation correlate between different educational subgroups. The Y-axis represents the size of school fixed effects for higher educated parents and X-axis displays the fixed effects for middle educated. The slope of the fitted values line is .932 and statistically different from zero at the $p < 0.01$ level. The slope of the line is not statistically different from one which means that we cannot reject that middle and higher educated have the same revealed preferences for schools.

Figure 5b: Correlation of school preferences between higher and lower educated parents



Note: The figure shows how school fixed effects obtained from conditional logit estimation correlate between different educational subgroups. The Y-axis represents the size of school fixed effects for higher educated parents and X-axis displays the fixed effects for middle educated. The slope of the fitted values line is .420 and statistically different from zero at the $p < 0.01$ level. The slope of the line is statistically different from one at the $p < 0.01$ level..

Figure 5c: Correlation of school preferences between middle and lower educated parents



Note: The figure shows how school fixed effects obtained from conditional logit estimation correlate between different educational subgroups. The Y-axis represents the size of school fixed effects for higher educated parents and X-axis displays the fixed effects for middle educated. The slope of the fitted values line is .361 and statistically different from zero at the $p < 0.01$ level. The slope of the line is statistically different from one at the $p < 0.01$ level.

TABLES

Table 1: Descriptive statistics

Type of school	(1) Public		(2) Catholic		(3) Protestant		(4) Islamic		(5) Alternative		(6) All schools	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Percentage of parents lower educated	0.30	0.15	0.27	0.16	0.33	0.22	0.59	-	0.23	0.14	0.16	0.00
Percentage of parents middle educated	0.30	0.09	0.37	0.11	0.32	0.16	0.20	-	0.35	0.13	0.35	0.11
Percentage of parents higher educated	0.40	0.16	0.37	0.16	0.34	0.17	0.22	-	0.41	0.21	0.37	0.16
Percentage with “weak” Inspectorate evaluation	0.10	0.30	0.08	0.28	0.11	0.33	0.00	-	0.14	0.36	0.09	0.29
Std. school average CITO	0.02	0.76	0.08	0.97	-0.12	1.05	2.93	-	-0.68	0.95	0.03	0.98
Student distance in kilometers	0.993	1.13	0.713	0.77	0.854	0.98	2.628	3.10	1.457	1.87	0.83	1.01
Distance rank of the school	3.438	4.56	2.094	3.21	3.475	4.64	11.16	11.02	5.781	7.07	2.66	4.15
Number of schools	31		145		9		1		14		200	
Average number of students	264.2		193.7		56.6		102.0		71.9		189.5	

Notes: The school type “Alternative” refers to schools with a particular teaching concept. Schools in this category are Montessori, Jenaplan and “vrije” schools.

Table 2: Number of schools in the neighborhood

Distance radius from home	Number of schools			
	Mean	SD	Min	Max
500 meters	0.84	0.76	0	4
1 km	2.24	1.39	0	9
2 km	6.23	3.69	0	17
3 km	11.66	5.90	0	27
4 km	17.74	7.86	0	34
5 km	24.62	9.52	0	40
Number of observations	15435			

Table 3: Preferences for school distance and quality - Estimates from conditional logit models

	(1)	(2)	(3)	(4)
	Conditional logit	Conditional logit	Conditional logit	Conditional logit
Distance to school (100m)	-0.211*** (0.015)	-0.211*** (0.015)	-0.211*** (0.015)	-0.211*** (0.015)
‘Weak’ Inspectorate evaluation		-0.401** (0.157)		-0.283* (0.165)
Std. CITO score			0.144*** (0.047)	0.127*** (0.049)
Observations	554,855	554,855	554,855	554,855
Number of schools	183	183	183	183
Number of students	15000	15000	15000	15000

Notes: The variable “std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable "weak" Inspectorate assessment includes the one school that was rated as “unsatisfactory”. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 4: Preferences for school distance and quality - Estimates from mixed logit models

	(1) Mixed logit	(2) Mixed logit	(3) Mixed logit	(4) Mixed logit
Means				
Distance to school (100m)	-0.308*** (0.005)	-0.309*** (0.005)	-0.312*** (0.005)	-0.313*** (0.005)
'Weak' evaluation		-0.497*** (0.049)		-0.336*** (0.050)
Std. CITO score			0.190*** (0.012)	0.169*** (0.012)
Standard deviations				
SD distance to school (100m)	0.131*** (0.003)	0.132*** (0.003)	0.134*** (0.003)	0.134*** (0.003)
SD 'weak' evaluation		0.038 (0.043)		0.033 (0.044)
SD std. CITO score			-0.013 (0.030)	0.030 (0.032)
Observations	554,855	554,855	554,855	554,855
Number of schools	183	183	183	183
Number of students	15000	15000	15000	15000

Notes: The variable “Std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable "weak" Inspectorate assessment includes the one school that was rated as “unsatisfactory”. For the estimations we use simulated maximum likelihood and specify the mixing distribution of all random variables to be normally distributed. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5: Heterogeneous preferences for school quality - Estimates from conditional logit models

	(1) Higher educated	(2) Middle educated	(3) Lower educated
Distance to school (100m)	-0.216*** (0.021)	-0.226*** (0.018)	-0.212*** (0.019)
Std. CITO score	0.256*** (0.085)	0.086 (0.061)	0.084 (0.061)
'Weak' Inspectorate evaluation	-0.811*** (0.294)	-0.448** (0.227)	-0.269 (0.205)
Observations	136,460	124,750	88,735
Number of schools	183	183	183
Number of students	3687	3365	2399

Notes: The variable “std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable "weak" Inspectorate assessment includes the one school that was rated as “unsatisfactory”. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 6: Heterogeneous preferences for school quality - Estimates from mixed logit models

	(1) Higher educated	(2) Middle educated	(3) Lower educated
Means			
Distance to school (100m)	-0.294*** (0.007)	-0.347*** (0.010)	-0.324*** (0.011)
Std. CITO score	0.287*** (0.026)	0.121*** (0.030)	0.144*** (0.033)
'Weak' Inspectorate evaluation	-1.037*** (0.228)	-0.517*** (0.115)	-0.316** (0.125)
Standard deviations			
SD Distance to school (100m)	0.112*** (0.006)	0.149*** (0.007)	0.145*** (0.008)
SD 'weak' evaluation	-0.015 (0.073)	0.018 (0.066)	-0.300*** (0.090)
SD std. CITO score	-0.723 (0.462)	0.091 (0.445)	0.130 (0.365)
Observations	136,460	124,750	88,735
Number of schools	183	183	183
Number of students	3687	3365	2399

Notes: The variable “Std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable "weak" Inspectorate assessment includes the one school that was rated as “unsatisfactory”. For the estimations we use simulated maximum likelihood and specify the mixing distribution of all random variables to be normally distributed. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 7: Heterogeneous preferences for school quality - Estimates from mixed logit models

	(1)	(2)	(3)	(4)
	All	Higher educated	Middle educated	Lower educated
Mean				
Distance to school (100m)	-0.337*** (0.007)	-0.325*** (0.010)	-0.375*** (0.013)	-0.336*** (0.013)
Std. CITO score	0.253*** (0.021)	0.411*** (0.038)	0.154*** (0.036)	0.146*** (0.030)
'Weak' Inspectorate evaluation	-0.458*** (0.085)	-1.057*** (0.306)	-0.389*** (0.143)	-0.311** (0.124)
Public school	-0.135** (0.062)	-0.013 (0.095)	-0.460*** (0.116)	0.107 (0.106)
Protestant school	-0.731*** (0.201)	-1.053** (0.508)	-0.973* (0.527)	-0.628*** (0.141)
Islamic school	-16.955*** (5.002)	-7.950** (3.323)	-26.946* (15.459)	-39.870*** (8.118)
Alternative teaching philosophy	-0.262** (0.126)	-1.056*** (0.351)	-0.260 (0.174)	-0.473 (0.395)
Standard deviations				
SD Distance to school (100m)	0.141*** (0.005)	0.130*** (0.007)	0.162*** (0.009)	0.149*** (0.016)
SD std. CITO score	-0.076 (0.050)	-0.010 (0.211)	0.029 (0.042)	-0.311 (0.205)
SD 'Weak' Inspectorate evaluation	-0.238 (0.145)	-0.920* (0.520)	0.130 (1.557)	0.072 (0.163)
SD Public school	-1.632*** (0.198)	1.872*** (0.281)	1.906*** (0.309)	-1.079 (0.770)
SD Protestant school	0.520 (0.501)	-1.023 (0.643)	1.194 (0.905)	-0.097 (0.346)
SD Islamic school	-13.236*** (3.869)	-5.438** (2.555)	19.227* (10.288)	-30.904*** (5.523)
SD Alternative teaching philosophy	0.237 (0.452)	2.095*** (0.432)	0.073 (0.342)	0.426 (0.439)
Observations	349,945	136,460	124,750	88,735
Number of schools	183	183	183	183
Number of students	9451	3687	3365	2399

Notes: The variable “Std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable "weak" Inspectorate assessment includes the one school that was rated as “unsatisfactory”. For the estimations we use simulated maximum likelihood and specify the mixing distribution of all random variables to be normally distributed. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 8: Preferences for school types by religion of parents' - Estimates from mixed logit models

	(1) Non-religious	(2) Catholic	(3) Islamic	(4) Others
Mean				
Distance to school (100m)	-0.352*** (0.020)	-0.358*** (0.010)	-0.363*** (0.044)	-0.305*** (0.016)
Std. CITO score	0.363*** (0.056)	0.202*** (0.026)	0.190 (0.131)	0.301*** (0.051)
'Weak' Inspectorate evaluation	-0.782** (0.333)	-0.584*** (0.101)	-0.174 (0.345)	-0.081 (0.195)
Public school	1.063*** (0.147)	-0.599*** (0.101)	0.395 (0.285)	0.301** (0.121)
Protestant school	-1.095 (0.693)	-1.241*** (0.184)	-0.648 (0.652)	0.334 (0.266)
Islamic school	-21.887*** (0.686)	-23.072*** (0.271)	5.239** (2.304)	-11.459** (5.478)
Alternative teaching philosophy	-0.467* (0.280)	-0.409*** (0.158)	-0.514 (0.927)	-0.264 (0.336)
Standard deviations				
SD Distance to school (100m)	0.158*** (0.014)	0.149*** (0.006)	0.147*** (0.026)	0.133*** (0.011)
SD std. CITO score	0.219 (0.179)	0.040 (0.050)	-0.267 (0.468)	-0.339** (0.164)
SD 'Weak' Inspectorate evaluation	0.965 (0.595)	0.191 (0.126)	-0.033 (0.150)	-0.322 (0.646)
SD Public school	-1.964*** (0.344)	1.308*** (0.338)	-0.965 (0.923)	1.806*** (0.365)
SD Protestant school	-1.217 (0.854)	-0.019 (0.426)	0.262 (0.920)	0.158 (3.634)
SD Islamic school	0.097 (0.076)	0.048* (0.026)	16.184*** (3.560)	-7.747*** (2.720)
SD Alternative teaching philosophy	-1.122*** (0.411)	0.582 (0.367)	0.766 (2.287)	0.854 (0.740)
Observations	55,661	223,757	10,869	59,658
Number of schools	183	183	183	183
Number of students	1510	6028	295	1618

Notes: The variable “Std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable "weak" Inspectorate assessment includes the one school that was rated as “unsatisfactory”. For the estimations we use simulated maximum likelihood and specify the mixing distribution of all random variables to be normally distributed. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

APPENDIX

Table A1: School level predictors of survey response – Probit regression

	(1) Response
Distance to school (100m)	-0.003** (0.001)
Std. CITO score	0.056*** (0.012)
'Weak' inspectorate evaluation	-0.133*** (0.046)
Catholic school	reference
Public school	-0.133*** (0.029)
Protestant school	-0.381*** (0.073)
Islamic school	-0.668*** (0.131)
Alternative teaching philosophy	0.070 (0.050)
Constant	0.454*** (0.015)
Observations	15,000
Number of schools	183

Notes: The variable “Std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable “‘weak’ Inspectorate assessment” includes the one school that was rated as “unsatisfactory”. Standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2: Differences between the full and the questionnaire sample

Sample:	Full	Questionnaire	Full	Questionnaire
	(1)	(2)	(3)	(4)
	Conditional logit	Conditional logit	Mixed logit	Mixed logit
Distance to school (100m)	-0.211*** (0.015)	-0.218*** (0.016)	-0.313*** (0.005)	-0.315*** (0.006)
'Weak' inspectorate evaluation	-0.283* (0.165)	-0.429** (0.207)	-0.336*** (0.050)	-0.479*** (0.080)
Std. CITO score	0.127*** (0.049)	0.150*** (0.058)	0.169*** (0.012)	0.190*** (0.016)
Standard deviations				
Distance to school (100m)			0.134*** (0.003)	0.130*** (0.004)
Std. SAT score			0.033 (0.044)	0.082 (0.086)
'Weak' inspection valuation			0.030 (0.032)	-0.012 (0.020)
Observations	554,855	362,019	554,855	362,019
Number of schools	183	183	183	183
Number of students	15000	9779	15000	9779

Notes: The variable “Std. CITO score” represents the standardized three year CITO score average from 2008, 2009 and 2010 at the school level. The variable “Weak Inspectorate assessment” includes the one school that was rated as “unsatisfactory”. For the estimations we specify the mixing distribution of all random variables to be normally distributed. The displayed coefficients represent the raw coefficients. Standard errors clustered at the school level are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1